

Aides à l'extraction des connaissances

- **Difficultés :**

- faire expliciter, capturer des connaissances explicites
 - interpréter sans introduire de biais

- **Travaux de recherche :**

- en psychologie, ergonomie et IA
 - établir des liens entre types de tâches et techniques d'extraction

- **Ce qui existe :**

- des techniques de recueil
 - des outils automatisant des techniques : ETS, AQUINAS, PC_PACK (Univ. de Nottingham) en amont de CommonKADS
 - des méthodes : MACAO, KOD

Techniques de recueil : caractéristiques

- **Place par rapport à l'activité** : préalable / concomitante / consécutive
- **Mode de communication** : direct, téléphone, courrier
- **Analyse des réponses** : qualitative ou quantitative
- **Types de connaissances recueillies** :
 - automatismes / habiletés : observations, tech. indirectes, expérimentations
 - règles : simulations dans le contexte, verbalisations
 - schémas, connaissances déclaratives: entretiens centrés
 - plans et stratégies : entretiens, simulations avec variations d'énoncés
- **Origine des données verbales** : spontanée (activité) / provoquée (questionnaires)
- **Situation du recueil** : conditions habituelles / transformées : simulations, entretiens ; en groupe / individuel
- **Documentation et compte rendu** : notes, retranscriptions
- **Mode d'accès aux connaissances** : direct (observations) / indirect (indicateurs)

Principes de recueil des données

● **Analyse des données recueillies**

- Enlever le « bruit », ce qui est « inutile »
- Identifier et extraire des fragments (concepts, relations, définitions)

Plus ou moins facile selon ce qui a été fait avant

Exemples : KOD (Vogel) , COGNOSYS (Woodward)

● **Choix des techniques**

- selon les psychologues : disposer d'une théorie de l'activité étudiée
- au moins : prendre en compte la diversité
 - des formes de connaissances
 - des niveaux d'abstraction et de régulation
 - des individus

Le modèle d'expertise : Qu'est ce qui est réellement modélisé ?

Jusque vers 1990 : KADS I, ...

modèle du monde réel en tant que modèle du processus de résolution de problème tel qu'il est *observé* à l'aide des experts du domaine ou d'autres sources

Après 1990 : KADS II, ...

spécification fonctionnelle de la partie résolution de problème de l'*artefact* à *construire*

On passe de l'observation à l'interprétation

Qu'est-ce qu'un modèle ?

● Rôles d'un modèle conceptuel :

- ☞ *conceptualiser entités du domaine et interactions*
- ☞ *intermédiaire entre connaissances et code : spécification ?*
 - Langage partagé par l'expert, le cognitifien
 - Langage d'expression des connaissances
 - Langage « compréhensible » par l'artefact

Un modèle est une abstraction qui permet de réduire la complexité en se focalisant sur certains aspects, en fonction de certains buts

MAIS un modèle devrait permettre plus : manipuler les objets et interpréter les résultats de la manipulation

Le modèle conceptuel

- **Connaissances de raisonnement au sens large** : expriment comment une tâche va être effectuée. Utilisent des connaissances du domaine
- **Connaissances du domaine** : concepts manipulés et leurs relations.
« Ce qui est vrai » dans le domaine pour la tâche considérée
- **Modèle conceptuel** : modèle explicitant la manière dont le système résout ainsi que les connaissances qu'il utilise

Modèle conceptuel

=

Connaissances de raisonnement + connaissances du domaine

Le « *knowledge level* » de Newell - 1982 -

- **Nécessité d'un niveau au dessus du niveau symbolique**

But = décrire un système informatique

différencier connaissances et représentations

description du *comportement observé* du système à l'aide de connaissances, indépendamment de leur formalisation

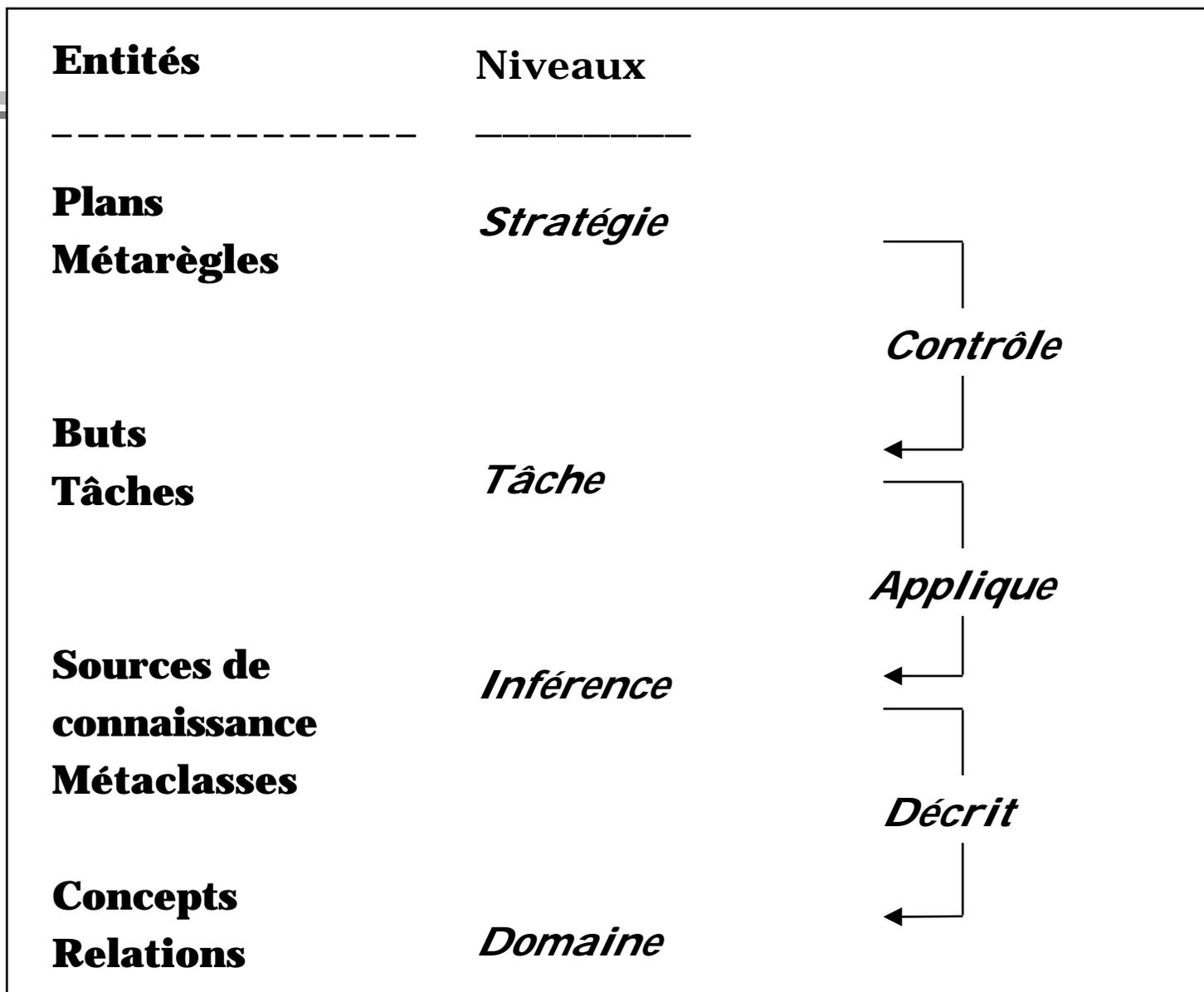
- **Agent rationnel**

dispose de *connaissances*, doit atteindre des *buts*, sait effectuer des *actions*, et est *rationnel* : choisit (avec ses connaissances) l'action suivante qui va le mener le plus directement au but

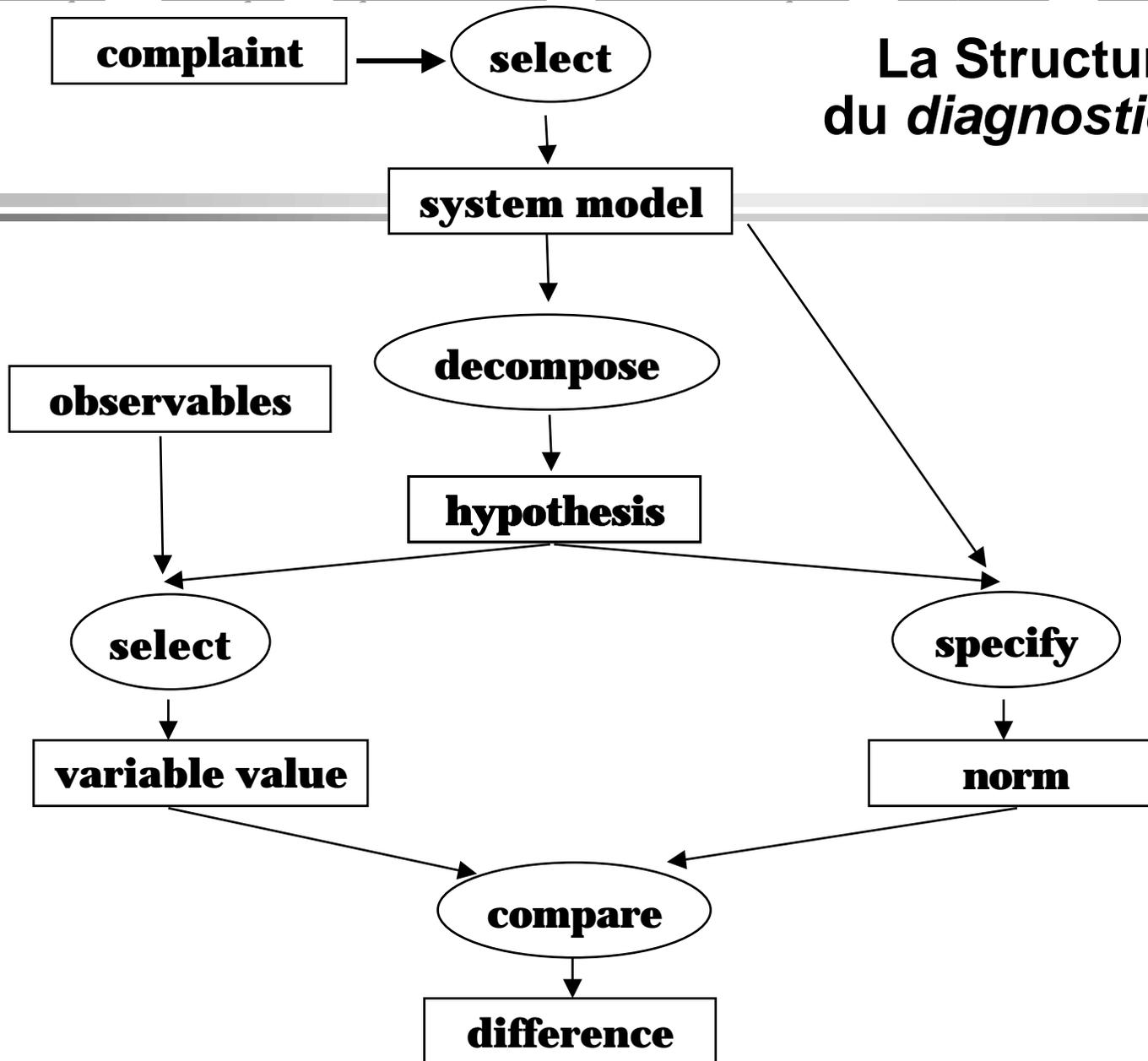
- **Application à la construction de SBC**

système = agent rationnel interagissant avec son environnement
descriptions plus structurées que ce que propose Newell.

KADS : Les quatre couches du modèle d'expertise



La Structure d'inférence du diagnostic systématique



Task Systematic-Diagnosis

Goal

find the smallest component with inconsistent behaviour if one

Control terms

differential = set of currently active hypothesis
 inconsistent-sub-system= subpart

Task-structure

systematic-diagnosis

(complaint —> inconsistent-sub-system) =
select

(complaint —> system-model)

generate-hypotheses

(system-model —> differential)

REPEAT

test-hypotheses

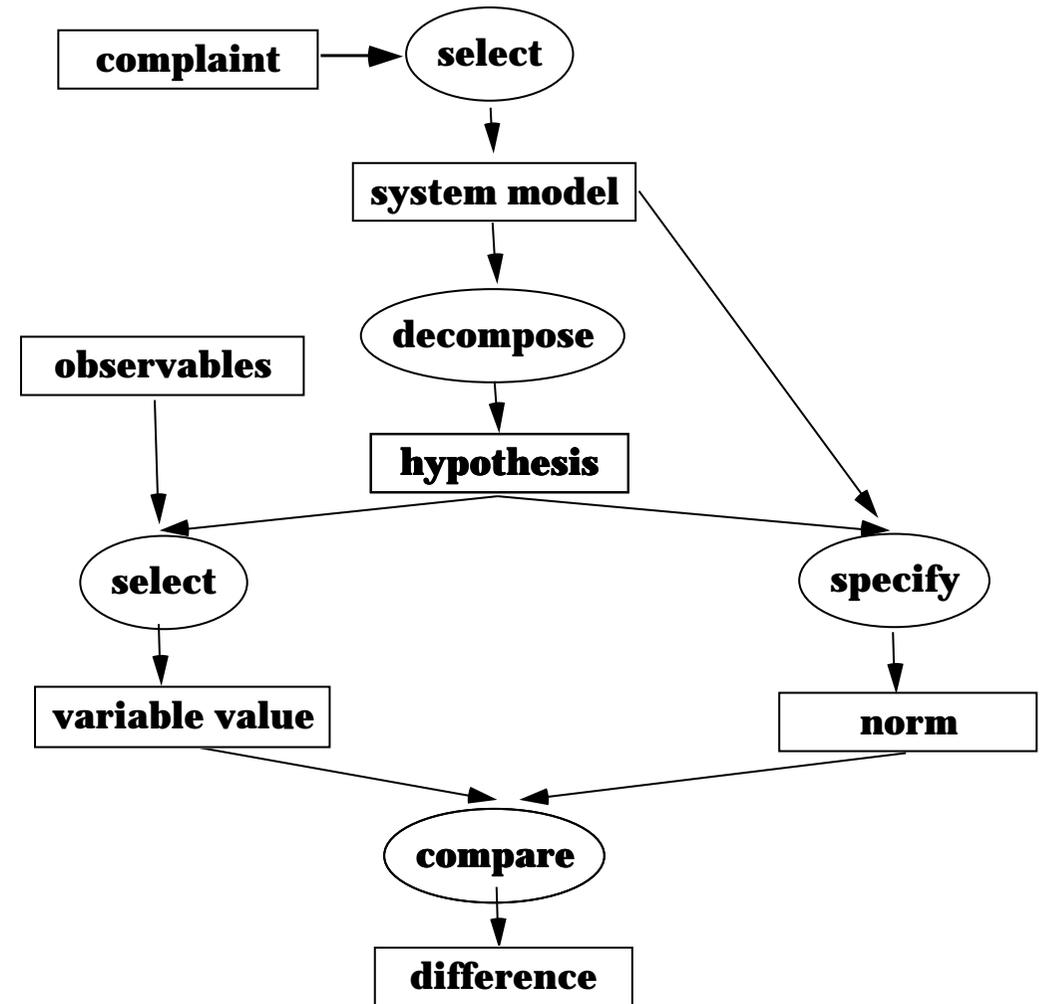
(differential —> inconsistent-sub-system)

generate-hypotheses

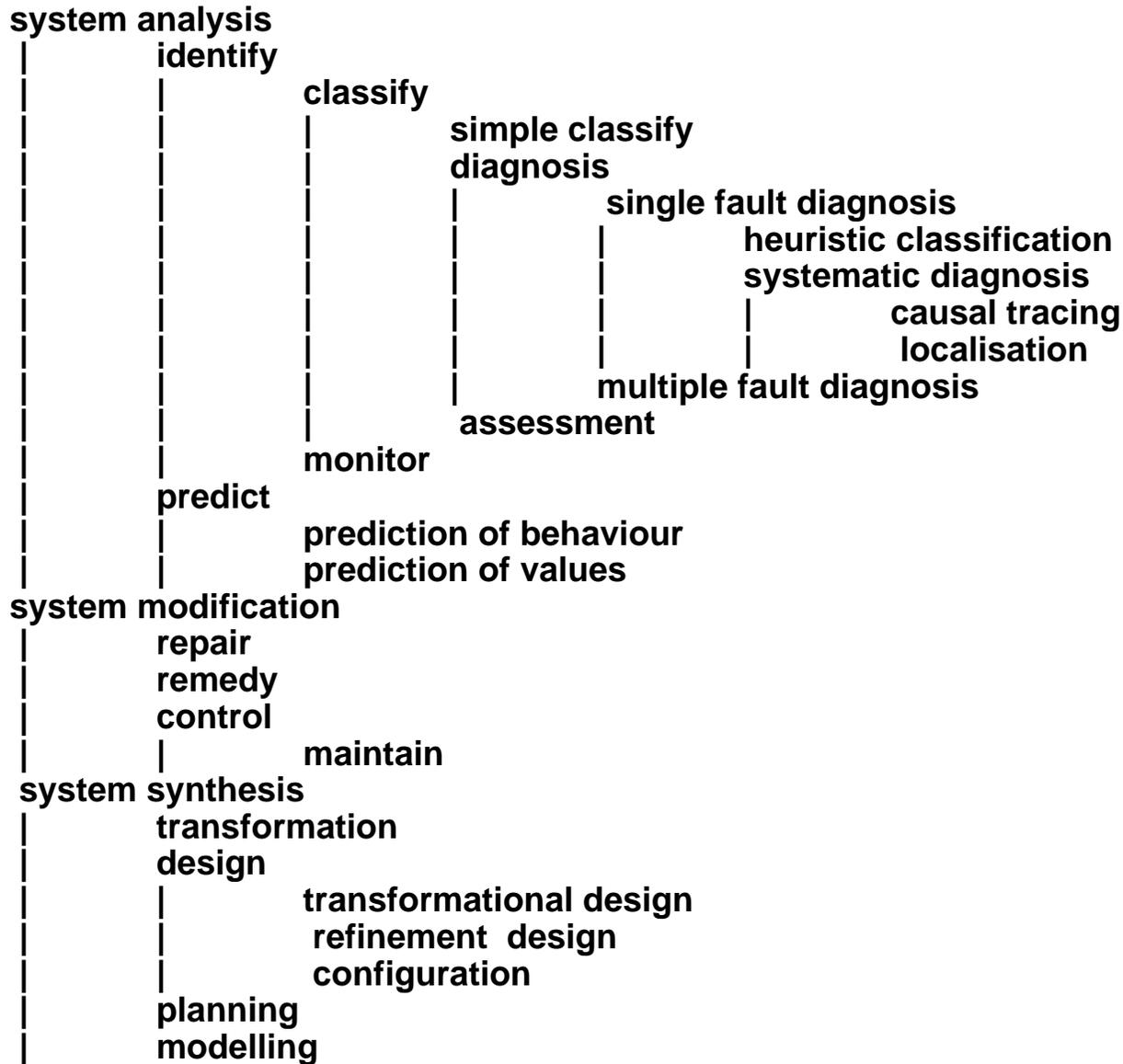
(inconsistent-sub-system —> differential)

UNTIL differential is empty

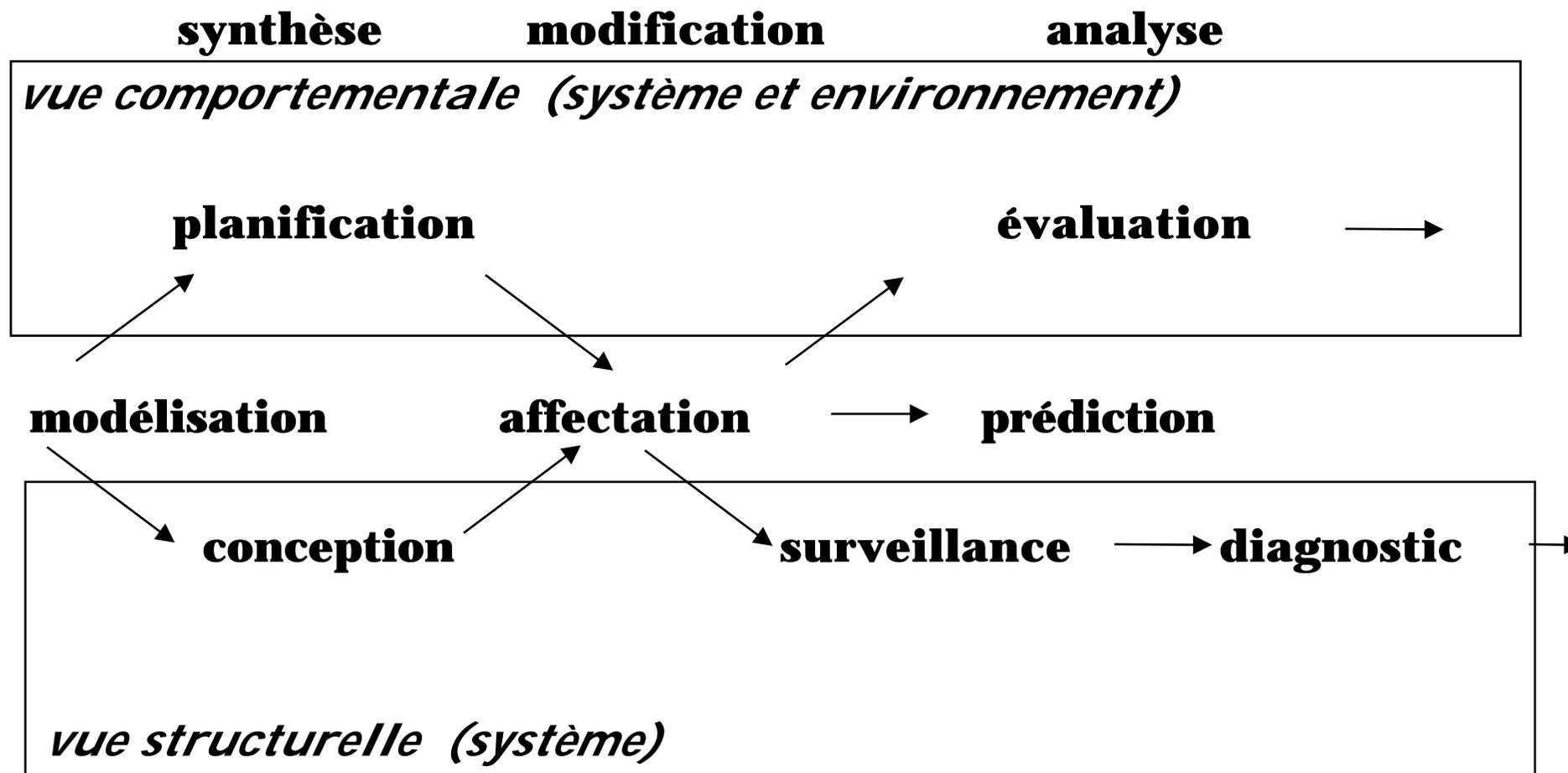
La Structure de tâche du diagnostic systématique (1)



KADS-1 : La bibliothèque des modèles d'interprétation



KADS-2 : imbrication des types de problèmes



Bibliothèques de composants réutilisables

- **Tâches : modèles et guides de modélisation**

Modèles papier et questions d'aide à la modélisation (livre CommonKADS – Chap. 6)

Modèles CML téléchargeables : WebCOKACE
<<http://www.inria.fr/acacia/Cokace>>

- **Méthodes de résolution de problèmes**

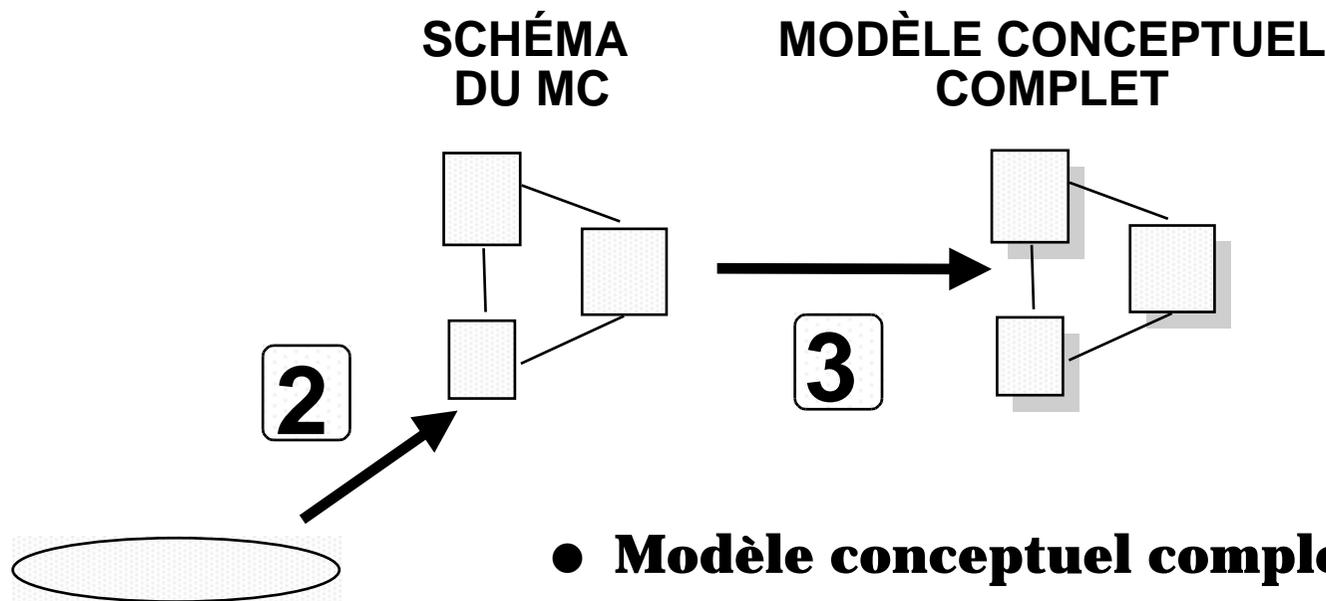
Inventaire de méthodes de résolution, d'opérateurs et de rôles applicables dans une méthode (livre CommonKADS – Chap. 13)

Modèles de méthodes en CML
<<http://www.swi.uva.nl/projects/CommonKADS/home.html>>

L'approche conduite par un modèle

- **Schéma du modèle conceptuel**

Modèle de raisonnement
+ Concepts du domaine



- **Modèle conceptuel complet**

Instanciación du schéma (instances de concepts,
inventaire de toutes les règles, etc.)

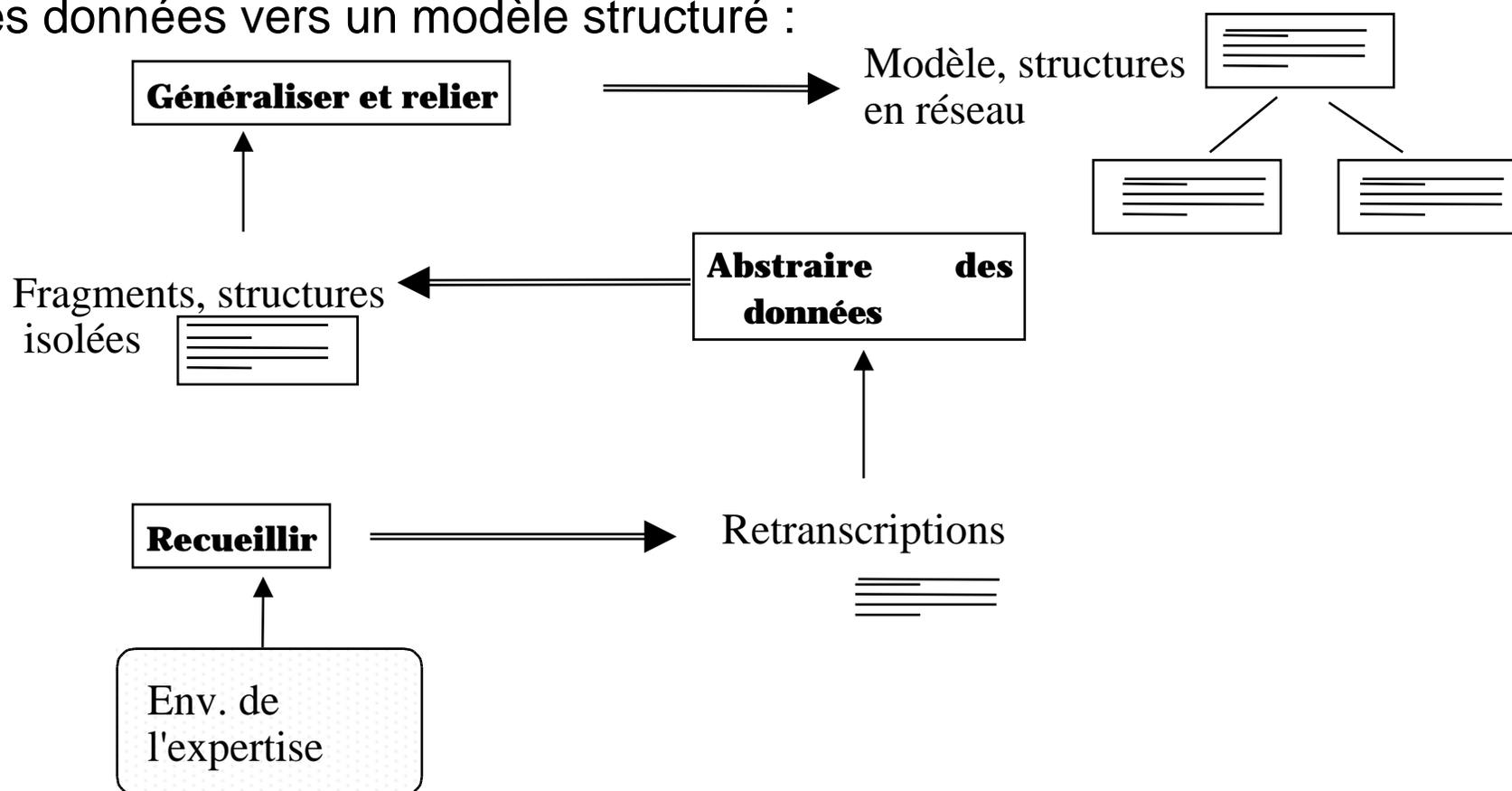
SADE : le projet

Lien vers description SADE

Construction ascendante *versus* descendante (1)

● Méthodes ascendantes :

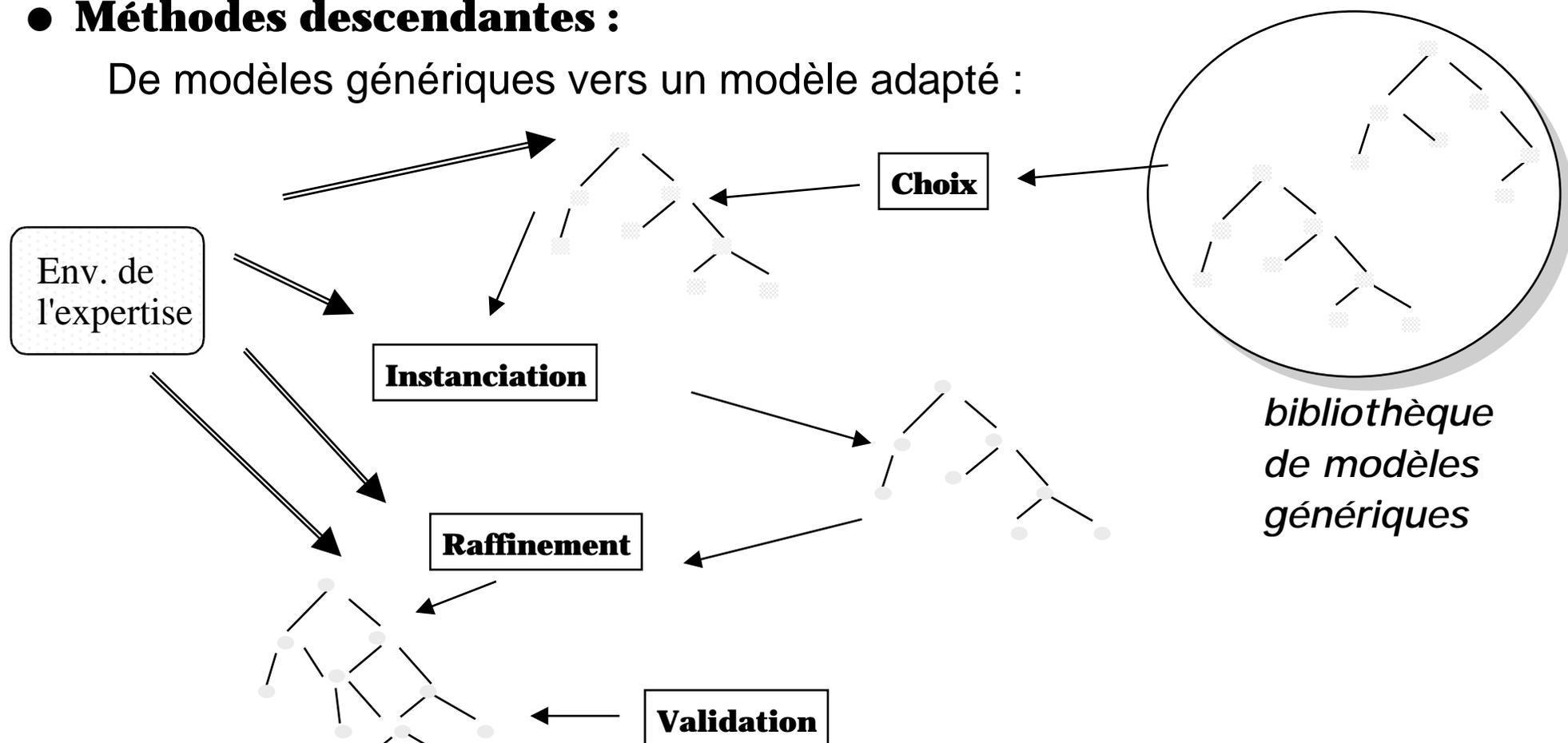
Des données vers un modèle structuré :



Construction ascendante *versus* descendante (2)

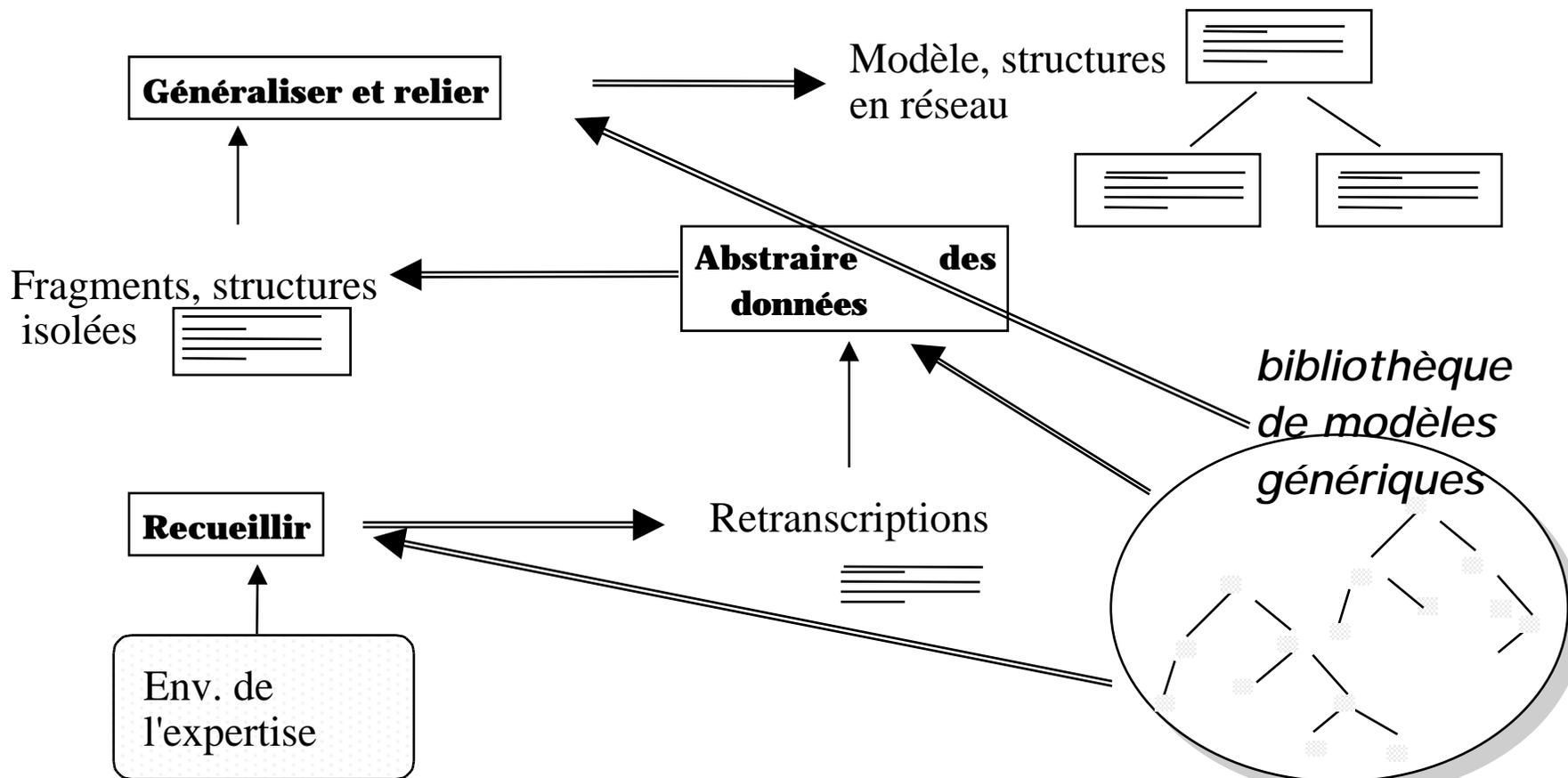
● Méthodes descendantes :

De modèles génériques vers un modèle adapté :



Construction ascendante versus descendante (3)

● Méthodes ascendantes : rôle des modèles génériques



Ascendant versus descendant : deux points de vue, deux cycles de base

Recueillir ... interroger, analyser, retranscrire

Conceptualiser ... abstraire, typer

Formaliser ... représenter

Structurer ... organiser

Valider

Caractériser ... typer, qualifier

Choisir ... comparer, réutiliser

Adapter ... modifier, corriger

Affiner ... préciser, compléter, spécifier

Instancier ... énumérer, lister

Valider

Exemple (1)

KADS

←→
complémentarité

MACAO

Caractérisation de la tâche

y consacrer peu de temps

Réutilisation

Comment choisir dans une bibliothèque ?

Comment répondre aux questions ?

Remarques

Une partie du travail d'abstraction est faite
au moment de construire la bibliothèque

Analyse initiale poussée

Catégorisation d'exemple : long

Simulations puis **modélisations** : long

Abstraction

Comment dégager le bon modèle ?

Comment mettre en évidence la
méthode ?

Remarques

Intérêt des catégories

Exemple (2)

KADS

Modèle du système

non opérationnel
quatre couches

Abstraction

Abstrait (en vue réutilisation)
Explicite bien une méthode de résolution
de problème

Utilisation

Peu accessible à l'expert (traduire)

MACAO

Modèle de l'expert

non opérationnel
deux parties : domaine / raisonnement

Abstraction

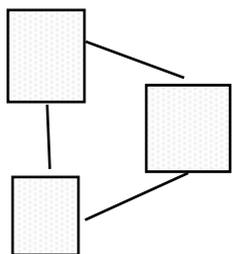
Proche du domaine
Méthode implicite sur graphe, visible
dans les structures

Utilisation

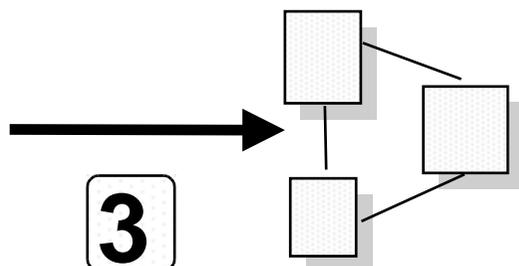
Lisible par l'expert mais pas assez
explicite
Niveau accessible = catégories

Instanciation du modèle conceptuel (1)

**SCHÉMA
DU MC**



**MODÈLE CONCEPTUEL
COMPLET**



● Objectifs

- Instancier le modèle conceptuel
- Affiner le schéma du modèle
- Être exhaustif

● Méthode

- Identifier tous les paramètres
- Expliciter les contrôles
- Définir les domaines de valeur
- Préciser règles, méthodes et procédures

*Processus de raffinement
progressif
Imbriqué avec étape 2*

Instanciación du modèle conceptuel (2)

- **On sait ce qu'il faut acquérir.**

Le modèle conceptuel spécifie précisément ce qu'il faut fournir au système. L'acquisition des connaissances est guidée par le modèle que l'on cherche à remplir. Chaque tâche, chaque méthode de résolution de problème nécessite des connaissances précises, pour un rôle bien spécifié.

- **Définition précise de la fonction des connaissances demandées.**

Une bonne définition et une bonne compréhension de la fonction (du rôle) des connaissances demandées en facilite l'expression par l'expert.

- **On dispose d'une sorte de métrique.**

On peut identifier facilement les endroits où il manque des connaissances.

- **Le système peut mieux « savoir » les limites de ce qu'« il sait ».**

Opérationnalisation du Modèle conceptuel (1)

● Objectifs

Traduire le MC en un modèle opérationnel

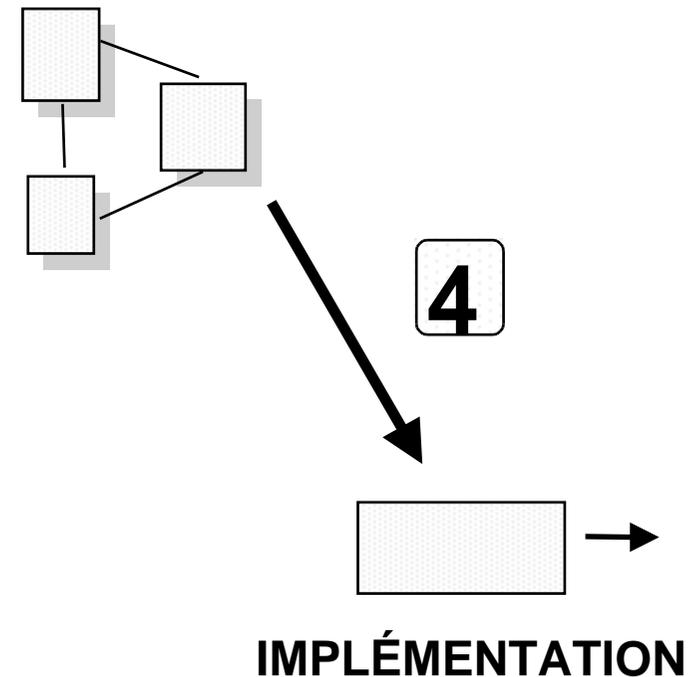
Valider le modèle par simulation

Développer la base de connaissances

● Contraintes

- correspondance BC ↔ MC
- conserver les liens entre structures
- anticiper les problèmes de maintenance

MODÈLE CONCEPTUEL COMPLET



Opérationnalisation : motivations

- **Motivations : Acquisition/validation et mise à jour**

- Conception incrémentale d'un modèle conceptuel

- Conception incrémentale d'une base de connaissances
(cycle acquisition, test, validation)

- Mise à jour de la base de connaissances
(les connaissances évoluent)

- **« Debugging » et maintenance :**

- Comportement observé dans le système opérationnel

- Compréhension / interprétation en termes du modèle conceptuel

- **Explications**

- en terme de connaissances

Opérationnalisation : solution

- **Créer des langages spécifiques pour représenter le modèle conceptuel**
 - Support pour l'acquisition des connaissances (MC, BC, ...)
- **Respecter le principe de correspondance structurelle entre modèle conceptuel et langage d'implémentation**
 - Possibilité de tracer le raisonnement en termes de buts et méthodes
 - Réflexivité permettant la mise en place d'un contrôle de bon niveau (ZOLA, DEF-*) ou à l'utilisateur d'interagir (LISA, ...)
- ...

Opérationnalisation : différentes approches (1)

- **Simuler/valider le modèle conceptuel (MODEL-K, ZOLA, OCML, OMOS,...)**
 - Possibilité de tracer le raisonnement en termes de *buts* et *méthodes*, etc.
 - Conception incrémentale d'un modèle conceptuel
 - Possibilité de validation fonctionnelle de partie du modèle (*tâche*, *méthode*
 - SACHEM)

Mais

- Représentation que de certains modèles (RLM, KADS – *tâche*, *inférence* [OMOS, MODEL-K] ; *tâche*, objet, classe d'objets, règles [AIDE])
- Perte de propriétés au moment de l'implémentation définitive

Opérationnalisation : différentes approches (3)

- **Coder directement le modèle opérationnel (LISA, KARL, DEF-*,)**

- Possibilité de tracer en permanence le raisonnement en termes de *buts* et *méthodes*
- Réflexivité permettant la mise en place d'un contrôle défini en termes de *buts*, *méthodes* permettant à l'utilisateur d'interagir (LISA, ...) ou la mise en place de primitives de représentation des connaissances manipulées par le langage (DEF-*)
- Possibilités de validation en introduisant des primitives spécifiques à cette validation (AMD)

Mais

- Pas toujours un langage réellement opérationnel (développement, gestion des erreurs, vitesse, ...)
- Représentation que de certains modèles (*buts*, *méthodes* [LISA], DEF-*)

Opérationnalisation : directions

- **Définir un métalangage, plus souple mais demandant plus de modélisation** (respect des modèles *tâche, méthode* mais possibilité de programmation « ad hoc » [ZOLA])
- **Développer une version « industrielle » du langage** (ZOLA IN JAVA, LISA@RT, DEF-*, ...)
- ...

De toute façon

- **Nécessité de faire cohabiter formes d'implémentation, de la plus déclarative à la plus procédurale**

Ateliers logiciels de modélisation

- **MOKA**

Projet européen, avec PSA, Daimler, BAES, Decan-Ingénia (développe et diffuse le logiciel), ... langage de modélisation spécifique, MML, extension d'UML

- **PROTEGE 2000**

Stanford Medical Informatics, Knowledge Modelling Group. Méta-langage de modélisation ; permet de décrire toute structure de modélisation et de générer des interfaces de saisie de connaissances pour ces structures (ex : CML - CommonKADS)

<<http://www-smi.stanford.edu/projects/protege/protege-2000/index.html>>